

LINGUISTIC TO LOGIC – THE SQL CONVERTER (ABSTRACT)

Introduction:

In this paper, we present a system for translating English statements into SQL queries, aiming to bridge the gap between users and database systems. This includes both non-technical users and technical users with limited knowledge to easily access database systems without explicitly writing code to access the data. We also propose frameworks for displaying requested images and generating plots, summaries and comparisons if requested by the end user through the question. We have included support for 3 built-in databases (Real Estate, Student and Medical Records) for users who want to check the project immediately and a Custom database support for users who want to use these and get the SQL Queries for their own institutional databases. We have a well defined UI for generating the table structure for those users.

Methodology:

The methodology proposed here is to use pretrained models (google gemini 1.0 pro and 1.5 pro) using prompts and prompt engineering. Two types of prompting were used, which are In-context learning (telling the model how and how not to act) and k-shot prompting (providing k examples of the task to be performed). As a single prompt and response is not enough for our problem, we used multiple prompts which are

- (1) Generating a boolean value for the different types of data requests (general, image, plot, summary and comparison). The boolean value for plot indicates whether a plot is requested or not.
- (2) Processing each of the individual components separately if it is set to True in the first response.
- (3) Generating English statements regarding what is requested and what is being displayed.

Then, we used pandas, mysql and matplotlib to process the data that we receive and the website was made using streamlit for python.